

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION  
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
29 mars 2001 (29.03.2001)

PCT

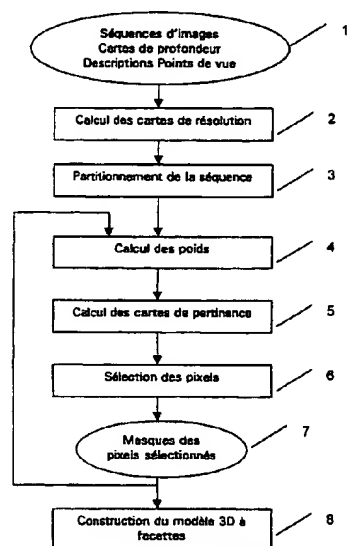
(10) Numéro de publication internationale  
WO 01/22366 A1

- (51) Classification internationale des brevets<sup>7</sup>: G06T 7/20 (72) Inventeurs; et  
(21) Numéro de la demande internationale: PCT/FR00/02537 (75) Inventeurs/Déposants (pour US seulement): ROBERT, Philippe [FR/FR]; Thomson multimedia, 46, quai Alphonse le Gallo, F-92648 Boulogne Cedex (FR). NICO-LAS, Yannick [FR/FR]; Thomson multimedia, 46, quai Alphonse le Gallo, F-92648 Boulogne Cedex (FR).  
(22) Date de dépôt international: 14 septembre 2000 (14.09.2000)  
(25) Langue de dépôt: français (74) Mandataire: BRAUN, Michel; Thomson multimedia, 46, quai Alphonse le Gallo, F-92648 Boulogne Cedex (FR).  
(26) Langue de publication: français (81) États désignés (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.  
(30) Données relatives à la priorité: 99/11671 17 septembre 1999 (17.09.1999) FR  
(71) Déposant (pour tous les États désignés sauf US): THOMSON MULTIMEDIA [FR/FR]; 46, quai Alphonse le Gallo, F-92100 Boulogne Billancourt (FR).

[Suite sur la page suivante]

(54) Title: METHOD FOR BUILDING A THREE-DIMENSIONAL SCENE BY ANALYSING A SEQUENCE OF IMAGES

(54) Titre: PROCEDE DE CONSTRUCTION D'UN MODELE DE SCENE 3D PAR ANALYSE DE SEQUENCE D'IMAGES



1...SEQUENCE OF IMAGES...DEPTH MAPS...DESCRIPTION OF VIEWPOINTS  
2...COMPUTING RESOLUTION MAPS  
3...PARTITIONING SEQUENCES  
4...COMPUTING WEIGHTS  
5...COMPUTING RELEVANCE MAPS  
6...SELECTING PIXELS  
7...MASKS FOR SELECTED PIXELS  
8...CONSTRUCTING THREE-DIMENSIONAL MODEL WITH FACETS

(57) Abstract: The invention concerns a method comprising the following steps: calculating, for an image, a depth map (1) corresponding to the depth, in the three-dimensional space, of image pixels; calculating, for an image, a resolution map (2) corresponding to the three-dimensional resolution of the image pixels, based on the depth map; matching (6) a current image pixel with a pixel of another image of the sequence, pixels relating to a common point in the three-dimensional scene, by projecting the current image pixel onto the other image; selecting a current image pixel (6) according to its resolution and to that of the other pixels of the sequence matched with said pixel; constructing the three-dimensional model (8) from the selected pixels. One application concerns the generation of images for browsing.

(57) Abrégé: Le procédé comporte les étapes suivantes: calcul, pour une image, d'une carte de profondeur (1) correspondant à la profondeur, dans l'espace 3D, des pixels de l'image; calcul, pour une image, d'une carte de résolution (2) correspondant à la résolution 3D des pixels de l'image, à partir de la carte de profondeur; mise en correspondance (6) d'un pixel d'une image courante avec un pixel d'une autre image de la séquence, pixels relatifs à un même point de la scène 3D, par projection du pixel de l'image courante sur l'autre image; sélection d'un pixel de l'image courante (6) en fonction de sa résolution et de celle des pixels d'autres images de la séquence mis en correspondance avec ce pixel; construction du modèle 3D (8) à partir des pixels sélectionnés. Une application concerne la génération d'images pour la navigation.

WO 01/22366 A1



(84) États désignés (régional): brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

**Publiée:**

— Avec rapport de recherche internationale.

Procédé de construction d'un modèle de scène 3D par analyse de  
séquence d'images.

5 L'invention concerne un procédé de construction d'un modèle de  
scène 3D par analyse de séquence d'images.

Le domaine est celui du traitement des séquences d'images et de  
la modélisation des scènes réelles statiques dans un contexte de navigation.  
La séquence est constituée d'images relatives à des scènes statiques à  
10 l'intérieur desquelles évolue le point de vue, c'est à dire la caméra.

L'objectif est de permettre à un utilisateur de naviguer virtuellement  
dans une scène réelle. Or les données sur la scène sont constituées de  
séquences d'images, qui peuvent représenter une énorme quantité  
d'informations. Ces séquences doivent être traitées de façon à fournir une  
15 représentation compacte de la scène, utilisable de façon optimale pour la  
navigation, c'est-à-dire permettant un rendu interactif, avec une qualité  
d'images contrôlée. Le problème est d'obtenir un taux de compression élevé  
tout en évitant les techniques de type prédictif inter-images, non adaptées à la  
navigation.

20 Différentes représentations de scènes existent actuellement. Il est  
possible de distinguer principalement :

- les représentations à partir de modèles 3D, dans lesquelles la  
géométrie de la scène est généralement représentée sous la forme de  
facettes planes auxquelles sont associées des images de texture. Cette  
25 modélisation est très utilisée pour représenter des scènes synthétiques  
obtenues par des logiciels de type CAO (conception assistée par ordinateur).  
En revanche, elle est encore peu utilisée pour représenter des scènes réelles,  
car elle est complexe. Les méthodes actuelles utilisent peu d'images,  
généralement des photographies, et les représentations qui en résultent sont  
30 sommaires, manquent de réalisme.

- les représentations non 3D obtenues par exemple à partir du  
logiciel QuickTime VR (marque déposée de la société Apple). Les données de  
la scène sont acquises sous la forme de panoramiques, avec séquences  
d'images de transition pour passer d'un panoramique à un autre. Une telle  
35 représentation limite considérablement les possibilités de navigation dans la  
scène virtuelle.

L'invention a pour but de pallier les inconvénients précités. Elle a pour objet un procédé de construction d'un modèle de scène 3D par analyse de séquence d'images, chaque image correspondant à un point de vue défini par sa position et son orientation, caractérisé en ce qu'il comporte les étapes

5 suivantes :

- calcul, pour une image, d'une carte de profondeur (1) correspondant à la profondeur, dans l'espace 3D, des pixels de l'image,
- calcul, pour une image, d'une carte de résolution (2) correspondant à la résolution 3D des pixels de l'image, à partir de la carte de profondeur,
- 10 - mise en correspondance (6) d'un pixel d'une image courante avec un pixel d'une autre image de la séquence, pixels relatifs à un même point de la scène 3D, par projection du pixel de l'image courante sur l'autre image,
- sélection d'un pixel de l'image courante (6) en fonction de sa résolution et de celle des pixels d'autres images de la séquence mis en correspondance
- 15 avec ce pixel,
- construction du modèle 3D (8) à partir des pixels sélectionnés.

20 Selon un mode particulier de réalisation, le procédé est caractérisé en ce que, les pixels sélectionnés d'une image constituant une ou plusieurs régions, des poids sont calculés et attribués aux pixels de l'image en fonction de leur appartenance ou non aux régions et des caractéristiques géométriques des régions auxquelles ils appartiennent dans l'image et en ce qu'une nouvelle sélection des pixels est effectuée en fonction des valeurs de résolution et de poids affectées aux pixels.

25

30 Selon un mode particulier de réalisation, qui peut être combiné au précédent, le procédé est caractérisé en ce qu'un partitionnement des images de la séquence est effectué en identifiant, pour une image courante, les images dont les points de vue correspondants ont un champ d'observation possédant une intersection avec le champ d'observation relatif à l'image courante, pour former une liste d'images qui lui est associée, et en ce que les autres images de la séquence pour lesquelles la mise en correspondance des pixels de l'image courante est effectuée sont les images de sa liste.

35 Le partitionnement des images de la séquence peut être complété en enlevant de la liste associée à une image, les images qui possèdent trop peu de pixels correspondant à ceux de l'image courante.

L'invention concerne également un procédé de navigation dans une scène 3D consistant à créer des images en fonction du déplacement du point de vue, caractérisé en ce que les images sont créées à partir du procédé de construction du modèle 3D précédemment décrit.

5

Les séquences d'images représentent une quantité de données très importante avec une forte redondance inter-images. L'utilisation d'un modèle 3D qui est le meilleur modèle de représentation d'une scène réelle statique et la mise en correspondance des images par de simples transformations géométriques permettent d'identifier largement la redondance inter-images. Ce modèle permet en effet de prendre en compte un grand nombre d'images. D'autre part, il ne nécessite pas d'opérations de compensation de mouvement au niveau de l'image 2D.

Un meilleur compromis entre compacité, c'est à dire compression des données à stocker et traiter, interactivité et qualité de rendu est réalisé : malgré le taux de compression élevé, le procédé fournit des images de bonne qualité et permet une grande souplesse et rapidité dans la navigation.

L'invention permet d'obtenir un meilleur réalisme que celui obtenu avec les techniques de modélisation 3D actuelles ainsi qu'une meilleure flexibilité que celle obtenue avec les techniques classiques de codage d'images.

Les caractéristiques et avantages de la présente invention ressortiront mieux de la description suivante, donnée à titre d'exemple et en référence aux figures annexées où :

- la figure 1 représente un algorithme de traitement décrivant les étapes d'un procédé selon l'invention,
- la figure 2 représente les référentiels associés à un point de vue.

L'acquisition des données de la scène réelle est étroitement liée à la représentation envisagée. Dans notre exemple, nous nous plaçons dans la situation où les images sont acquises par une caméra standard, à la cadence vidéo, et le mouvement de caméra est élaboré de manière à correspondre aux trajets prévus en cours d'exploitation. Dans ce contexte, la construction d'une représentation de scène à partir de séquences d'images peut être apparentée aux techniques de codage d'images.

Le principe de construction de la représentation d'une scène est de sélectionner les données nécessaires et suffisantes assurant la reconstruction des images de la séquence avec une qualité contrôlée. Le processus consiste à comparer les images une à une pour sélectionner les régions ayant la  
5 meilleure pertinence, paramètre qui dépend de la résolution et du coût de description. En fait, la comparaison est effectuée au niveau des pixels : le critère de base pour la comparaison et la sélection des pixels est la résolution de la surface locale 3D associée à chaque pixel.

Nous supposons que par un traitement adéquat, connu de l'art  
10 antérieur, nous obtenons, pour chaque point de vue, sa position 3D dans un référentiel associé à la scène (position et orientation du point de vue), ainsi qu'une carte de profondeur associée à l'image relative au point de vue. La phase suivante a pour objet de constituer une représentation compacte de l'ensemble de ces données adaptée à la navigation.

15 La figure 1 représente un organigramme décrivant les différentes étapes du procédé selon l'invention.

Nous disposons en entrée du système, référencée 1, des données relatives à une séquence d'images acquise par une caméra se déplaçant dans  
20 une scène réelle statique comme indiqué précédemment. Il est cependant tout à fait envisageable que certains objets mobiles soient présents dans l'image. Dans ce cas, un traitement spécifique identifie ces objets qui sont alors marqués pour être ignorés lors des traitements ultérieurs. Un traitement ad hoc fournit, pour chaque image, une carte de profondeur ainsi que la position  
25 et l'orientation du point de vue correspondant. Il n'y a pas d'information de profondeur dans les zones correspondant à des objets mobiles supprimés.

Pour chaque pixel de chaque image est calculée une valeur de résolution, c'est l'étape 2. Un premier et un deuxième partitionnement sont ensuite réalisés lors de l'étape 3. L'étape 4 effectue un calcul de poids pour  
30 fournir, étape 5, des valeurs de pertinence attribuées aux pixels. L'étape suivante 6 réalise une sélection des pixels en fonction de leur pertinence. Une séquence de masques des pixels sélectionnés est ensuite obtenue pour la séquence d'images, à l'étape 7. Après cette étape 7, les étapes 4 à 7 sont  
35 répétées pour affiner les masques. Ces étapes sont répétées jusqu'à ce que les masques n'évoluent plus de manière significative. Alors, l'étape 8 est abordée pour réaliser la construction du modèle 3D à facettes à partir des seuls pixels sélectionnés.

Les différentes étapes sont maintenant expliquées en détail.

En entrée du système est disponible, pour chaque image de la séquence, une carte de profondeur ainsi que la position et l'orientation du point de vue correspondant.

L'étape 2 consiste en un calcul, pour chaque pixel d'une image, d'une valeur de résolution donnant une carte de résolution pour l'image.

La résolution en chaque pixel fournit une indication du niveau de détail de la surface telle qu'elle est vue du point de vue courant. Elle peut être, par exemple, calculée sur un bloc de points centré sur le pixel et correspond à la densité des points dans la scène, c'est à dire dans l'espace 3D, relatifs à ce bloc.

Dans un exemple, une fenêtre de 7x7 pixels centrée sur le pixel de l'image pour lequel la résolution est calculée, est exploitée. Pour chacun des pixels appartenant à cette fenêtre, l'information de profondeur est traitée pour déterminer, à partir de la répartition dans l'espace 3D des points autour du pixel traité, la résolution 3D : une répartition des points sur une grande profondeur donnera une moins bonne résolution qu'une répartition des points sur une faible profondeur. Après traitement de tous les pixels de l'image, on obtient une carte de résolution de l'image et cela pour chacune des images de la séquence.

Le procédé réalise ensuite, étape 3, une partition de la séquence.

La phase de navigation consiste à interpoler l'image du point de vue courant à partir du modèle 3D. Le modèle peut être très grand, et il est donc utile de le partitionner afin de limiter à chaque instant la quantité d'informations à traiter pour la reconstruction d'un point de vue. En effet, il est important que les images soient interpolées en temps limité afin de garantir une bonne fluidité de navigation. Par ailleurs, la comparaison des images pixel à pixel dans la phase 6 de sélection des données, décrite plus loin, est une opération lourde, en particulier si les séquences sont longues. Cette remarque plaide aussi pour un partitionnement, effectué le plus tôt possible, pour réduire la quantité de calculs.

Deux opérations de partitionnement sont en fait effectuées pour limiter la manipulation des données, tant dans la phase de construction de la représentation que dans la phase d'exploitation (navigation).

Un premier partitionnement de la séquence est effectué en identifiant les points de vue n'ayant pas d'intersection de leur champ

d'observation. Ce qui permettra d'éviter de les comparer, c'est à dire de comparer les images relatives à ces points de vue, lors des étapes suivantes. On détermine donc les intersections éventuelles entre les champs d'observation, de forme pyramidale, de chaque point de vue, en détectant les intersections entre les bords de ces champs. Cette opération ne dépend pas du contenu de la scène, mais uniquement de la position relative des points de vue. A chaque image courante est ainsi associé un ensemble d'images dont le champ d'observation possède une intersection avec celui de cette image courante, cet ensemble constituant une liste.

Une projection est effectuée lors de cette étape 3 de partitionnement permettant un deuxième partitionnement. Pour chaque groupe d'image, une projection, similaire à celle décrite plus loin en regard de l'étape 6, est réalisée pour identifier les pixels mis en correspondance. Si une image a trop peu de pixels mis en correspondance avec les pixels d'une image de sa liste, cette dernière est supprimée de la liste.

Il résulte de ces partitionnements, pour chaque point de vue, une liste ou groupe de points de vue ayant des points 3D en commun avec lui, et qui vont donc être comparés lors de la sélection des pixels pour réduire la redondance. Un tableau est constitué afin d'identifier, pour chaque image de la séquence, les images sélectionnées nécessaires à sa reconstruction.

Lors de la projection, les pixels n'ayant aucune correspondance sont marqués en forçant la valeur de résolution, par exemple à 1. Grâce à ce marquage particulier, on saura, lors de l'étape 6, qu'il n'est pas nécessaire de reprojeter ces pixels pour la recherche des pixels mis en correspondance. Cette opération de projection est en effet répétée à l'étape 6 afin d'éviter la mémorisation des informations relatives à ces correspondances, obtenues lors de l'étape 3, ces informations représentant un très grand nombre de données.

L'étape 4 consiste en un calcul de poids pour chacun des pixels d'une image. Ce paramètre est introduit afin de prendre en compte le coût des pixels conservés. En l'absence de contrainte supplémentaire sur la sélection des pixels, ceux-ci peuvent constituer des régions de tailles et de formes diverses, et le coût de description de ces régions peut être élevé. Pour éviter ce problème, un poids prenant en compte la classification des pixels dans l'environnement proche (pixel sélectionné ou pas) est associé à chaque pixel. Le poids peut être choisi de façon à pénaliser les régions de petite taille ou plus grossièrement, les images ayant peu de points sélectionnés. Dans ce cas, ce peut être une valeur par image, par exemple le pourcentage de points

sélectionnés. Il est également possible d'appliquer des filtres morphologiques sur le masque décrivant les régions de points sélectionnés afin de réduire la complexité de leur forme et donc réduire le coût de description.

Les critères qui peuvent être pris en compte pour le calcul de poids sont par exemple :

- la quantité des points sélectionnés dans l'image
- la taille des régions
- la compacité des régions (inversement proportionnelle au poids)
- la zone périphérique des régions pour tenir compte par exemple

des pointes à éliminer. Un filtre morphologique peut également être passé sur le masque avant le calcul du poids pour supprimer ces zones périphériques de faible surface.

A la première itération, les masques sont initialisés à la valeur 0, c'est à dire que tous les pixels sont sélectionnés par défaut. Les poids calculés, lors de ce premier passage à l'étape 4, sont donc à la valeur unité. Une variante consiste à choisir, comme poids pour tous les pixels de l'image, lors de cette première itération, le pourcentage de points de l'image n'ayant pas de correspondant dans les autres images par rapport au nombre de points de l'image. On favorise ainsi la conservation des images contenant le plus de pixels sans correspondant (voir étapes 5 et 6 pour la sélection des pixels).

Une valeur de pertinence combinant la résolution et le poids est déduite lors de l'étape 5. Elle peut par exemple être calculée ainsi :

$$\text{pertinence} = \text{résolution} \times (1 + \text{poids})$$

Une valeur est attribuée à chaque pixel pour fournir une carte de pertinence par image.

L'objectif est ici d'obtenir le maximum des points décrivant la scène sur un minimum d'images, les pixels étant sélectionnés (voir étape 6) en fonction de leur valeur de pertinence.

La sélection des pixels est l'objet de l'étape 6

Il s'agit ici d'une recherche, pour chaque pixel, du correspondant dans les autres points de vue, et d'une comparaison des valeurs de pertinence pour l'identification du pixel à meilleure pertinence.

Pour ce faire, une correspondance entre les pixels des différentes images est effectuée par transformation géométrique. La figure 2 décrit un repère image (O, u, v) correspondant à une image i, c'est à dire une image associée à un point de vue i, un référentiel (Oci, xci, yci, zci) lié au point de

vue  $i$  (par exemple  $Oci$  confondu avec la position du point de vue  $i$ ) et un référentiel absolu ( $Oa, xa, ya, za$ ).

Pour chaque point de vue  $i$ , on dispose de sa position et de son orientation dans le référentiel absolu. Chaque pixel  $(u,v)$  de l'image dispose  
5 d'une valeur de profondeur  $zci(u,v)$  définie dans le repère ( $Oci, xci, yci, zci$ ) associé au point de vue  $i$ .

On connaît la transformation géométrique qui permet de passer du repère image ( $O, u, v$ ) au repère ( $Oci, xci, yci, zci$ ) lié au point de vue, ainsi que les transformations géométriques permettant de passer de ce repère au  
10 repère absolu ( $Oa, xa, ya, za$ ) lié à la scène.

Ce sont ces transformations qui sont utilisées pour passer d'une image à une autre, c'est à dire pour mettre en correspondance les pixels d'une image aux pixels d'une autre image, comme indiqué ci-après.

Chaque pixel est le résultat de la projection d'un point de l'espace  
15 3D sur le plan image 2D du point de vue courant  $i$ . Partant d'un pixel de l'image  $i$  (dont la composante  $z$  est connue), qui correspond à un point quelconque de la scène, on peut déterminer son point de projection dans une image  $j$  par transformation géométrique connue. Si ce point de projection coïncide avec un pixel de l'image, il y a correspondance des pixels. Dans le  
20 cas contraire, ce point de projection 2D est associé au pixel le plus proche. On considère alors que ces 2 pixels (celui de départ et celui d'arrivée), qui sont relatifs à des points très proches d'une même surface dans la scène, sont en correspondance et que leurs caractéristiques peuvent être comparées.

La mise en correspondance des pixels d'une image est effectuée  
25 sur l'ensemble des images de sa liste, objet de la partition définie à l'étape 3. Chaque pixel est projeté sur chacune des autres images du groupe: il est mis en correspondance avec un pixel comme décrit ci-dessus. La valeur de pertinence est comparée et le pixel ayant la moins bonne pertinence est marqué. Le processus de comparaison du pixel aux pixels correspondants est  
30 stoppé dès que l'on trouve un correspondant qui a une meilleure pertinence.

Ces opérations permettent donc d'identifier et supprimer la redondance inter-images en ne conservant que les pixels à meilleure pertinence. Toutefois, tout en réduisant la redondance, il peut être intéressant de conserver plus d'une représentation d'une surface donnée pour éviter  
35 d'avoir à manipuler la représentation à résolution maximale pour créer des points de vue lointains. Il est donc avantageux d'introduire un seuil sur la comparaison des valeurs de résolution : si le rapport entre deux valeurs de

résolution dépasse ce seuil, aucun des pixels n'est marqué. Ainsi, chacun des pixels peut être utilisé en fonction de la résolution souhaitée, fine ou grossière.

Le marquage des pixels se fait en initialisant tout d'abord tous les pixels de tous les masques par exemple à la valeur binaire 1. Chaque pixel est comparé avec son correspondant, s'il existe, dans les autres points de vue associés lors des phases de partitionnement. Celui qui possède la pertinence la plus faible est marqué à 0, c'est-à-dire qu'il est rejeté. En conséquence, si aucun de ses correspondants n'a une plus grande pertinence que le pixel courant, celui-ci est sélectionné, puisqu'il conserve le marquage initial. Il en résulte donc, pour chaque image de la séquence, un masque ou image binaire, les pixels pour lesquels la valeur 1 est attribuée correspondant aux pixels sélectionnés.

L'étape 7 regroupe les masques relatifs à chacune des images constituant la séquence pour fournir la séquence de masques.

L'étape 7 est rebouclée sur l'étape 4 pour affiner les valeurs de pertinence calculées. A chaque itération, les poids et donc les valeurs de pertinence sont recalculés à partir des masques obtenus à l'itération précédente.

Les opérations de projections sont reproduites à chaque itération et concernent l'ensemble des pixels de l'image, des pixels non sélectionnés lors d'une précédente itération pouvant être sélectionnés du fait par exemple d'une diminution de la valeur de pertinence d'un pixel avec qui il est mis en correspondance. Cependant, les pixels n'ayant pas de correspondant dans les autres images ne sont pas projetés.

Pour réduire les calculs, il est possible, à chaque itération, de supprimer de la liste des images associée à une image courante, les images n'ayant plus aucun pixel à meilleure pertinence que le pixel correspondant dans l'image courante. La liste finale d'une image donnée contient ainsi les images nécessaires et suffisantes pour sa reconstruction.

Le processus itératif est stoppé après un nombre déterminé d'itérations ou bien lorsqu'il n'y a plus de changements significatifs dans les masques. Une fois ces masques définitifs obtenus, l'étape 8 succède à l'étape 7 et ceux-ci sont exploités dans la phase de construction du modèle 3D à facettes, la construction étant effectuée à partir des seuls pixels sélectionnés, définis par ces masques.

Les données relatives à ce modèle 3D à facettes sont composées d'une information de géométrie et d'une information de texture. Pour chaque

- région sélectionnée, définie par les masques, sa silhouette est polygonisée et la carte de profondeur correspondante est approximée par des triangles 3D. Les données de texture sélectionnées sont regroupées de façon à ne pas conserver les zones inutiles. Un modèle 3D peut facilement être constitué à
- 5 partir de l'ensemble de ces informations. La liste des images et donc des régions associées à chaque image peut aussi être avantageusement prise en compte dans la constitution du modèle 3D pour partitionner celui-ci. Ce partitionnement peut ensuite être exploité dans la phase de rendu pour limiter la quantité d'informations à traiter lors de la reconstruction d'images.
- 10 Le procédé de navigation dans la scène 3D qui consiste à créer des images en fonction du déplacement du point de vue, exploite l'ensemble de ces informations pour recréer les images.

## Revendications

- 1 Procédé de construction d'un modèle de scène 3D par analyse de
- 5 séquence d'images, chaque image correspondant à un point de vue défini par sa position et son orientation, caractérisé en ce qu'il comporte les étapes suivantes :
- calcul, pour une image, d'une carte de profondeur (1) correspondant à la profondeur, dans l'espace 3D, des pixels de l'image,
  - 10 - calcul, pour une image, d'une carte de résolution (2) correspondant à la résolution 3D des pixels de l'image, à partir de la carte de profondeur,
  - mise en correspondance (6) d'un pixel d'une image courante avec un pixel d'une autre image de la séquence, pixels relatifs à un même point de la scène 3D, par projection du pixel de l'image courante sur l'autre image,
  - 15 - sélection d'un pixel de l'image courante (6) en fonction de sa résolution et de celle des pixels d'autres images de la séquence mis en correspondance avec ce pixel,
  - construction du modèle 3D (8) à partir des pixels sélectionnés.
- 20 2 Procédé selon la revendication 1, caractérisé en ce que, les pixels sélectionnés d'une image constituant une ou plusieurs régions, des poids sont calculés et attribués aux pixels de l'image (4) en fonction de leur appartenance ou non aux régions et des caractéristiques géométriques des régions auxquelles ils appartiennent dans l'image et en ce qu'une nouvelle
- 25 sélection des pixels (6) est effectuée en fonction des valeurs de résolution et de poids affectées aux pixels.
- 3 Procédé selon la revendication 2, caractérisé en ce qu'une valeur de pertinence (5) est attribuée à chaque pixel d'une image en fonction du
- 30 poids et de la résolution affectés à ce pixel et en ce qu'une sélection des pixels (6) d'une image courante est effectuée à partir de la plus forte valeur de pertinence (5) parmi les pixels mis en correspondance pour donner un masque de pixels sélectionnés.
- 35 4 Procédé selon la revendication 1, caractérisé en ce qu'un partitionnement (2) des images de la séquence est effectué en identifiant, pour une image courante, les images dont les points de vue correspondants ont un

champ d'observation possédant une intersection avec le champ d'observation relatif à l'image courante, pour former une liste d'images qui lui est associée, et en ce que les autres images de la séquence pour lesquelles la mise en correspondance des pixels de l'image courante (6) est effectuée sont les  
5 images de sa liste.

5 Procédé selon la revendication 4, caractérisé en ce qu'un partitionnement (2) des images de la séquence est effectué en enlevant de la liste associée à une image, les images qui possèdent trop peu de pixels  
10 correspondant à ceux de l'image courante.

6 Procédé selon la revendication 3, caractérisé en ce que les opérations de calcul de poids (4), de calcul de pertinence (5) et de sélection des pixels sont réitérées jusqu'à ce que les masques obtenus à partir de la  
15 sélection n'évoluent plus de manière significative.

7 Procédé selon la revendication 3, caractérisé en ce que les opérations de mise en correspondance d'un pixel de l'image courante, par projection sur les autres images, sont arrêtées pour ce pixel dès qu'un pixel  
20 correspondant ayant une valeur de pertinence plus grande est trouvé.

8 Procédé selon la revendication 3, caractérisé en ce que la sélection à partir des valeurs de pertinence est effectuée lorsque le rapport entre les valeurs de résolution des pixels mis en correspondance se trouve à  
25 l'intérieur de limites prédéfinies.

9 Procédé selon la revendication 1, caractérisé en ce que le pixel de l'autre image est le pixel le plus proche du point de projection (6) sur cette autre image.  
30

10 Procédé selon la revendication 1, caractérisé en ce que les objets mobiles se déplaçant dans la scène sont détectés pour en être extraits afin d'obtenir une scène de type statique.

35 11 Procédé de navigation dans une scène 3D consistant à créer des images en fonction du déplacement du point de vue, caractérisé en ce

que les images sont créées à partir du procédé de construction du modèle 3D selon la revendication 1.

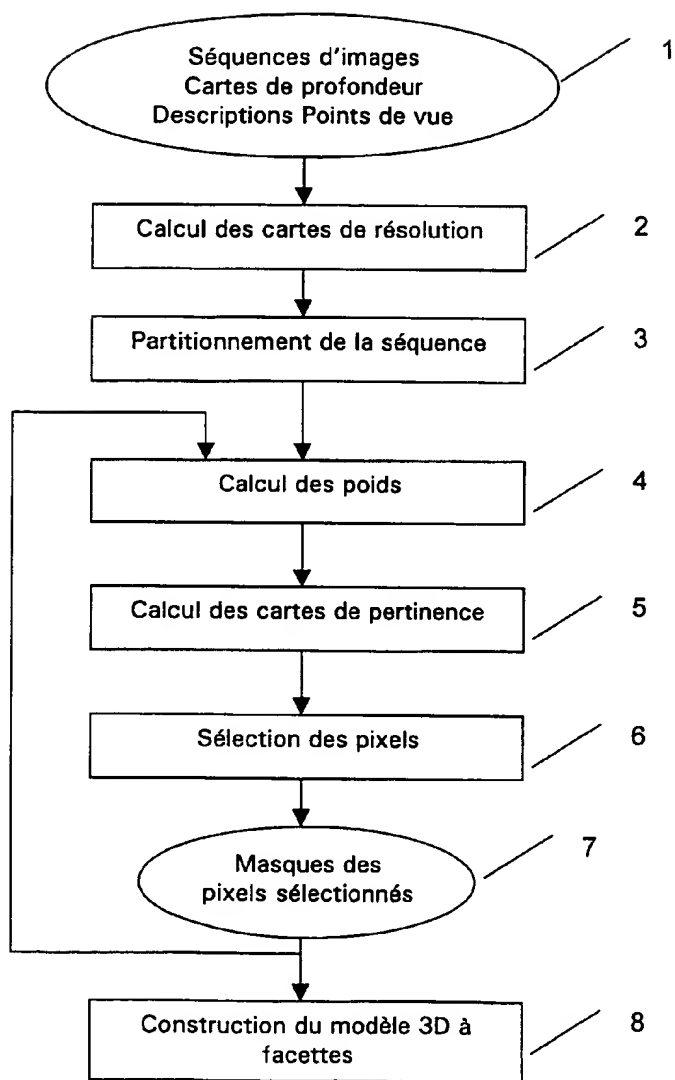


FIG.1

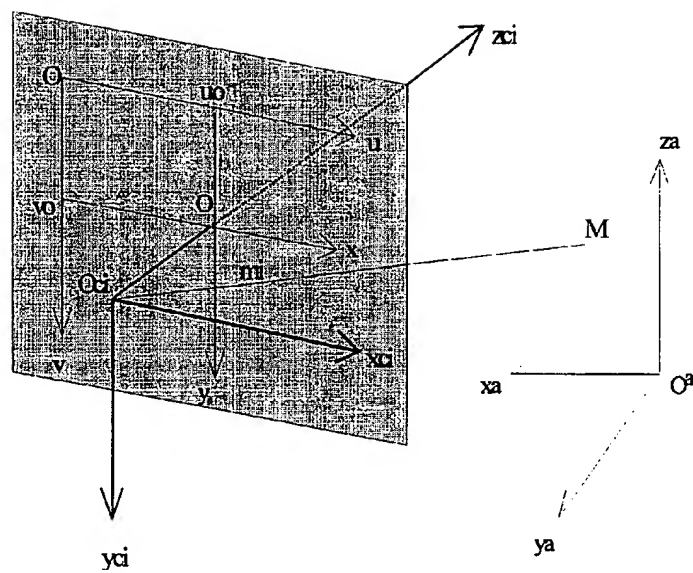


FIG.2

# INTERNATIONAL SEARCH REPORT

Int: onal Application No

PCT/FR 00/02537

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06T7/20

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, INSPEC, EPO-Internal, IBM-TDB

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 511 153 A (ALI AZARBAYEJANI ET AL.) 23 April 1996 (1996-04-23) abstract	1,11
A	EP 0 735 512 A (SANYO ELECTRIC CO., LTD.) 2 October 1996 (1996-10-02) abstract	1,11
Y	J.K. AGGARWAL ET AL.: "determining motion parameters using intensity guided range sensing" PATTERN RECOGNITION., vol. 19, no. 2, 1986, pages 169-180, XP002143346 PERGAMON PRESS INC. ELMSFORD, N.Y., US ISSN: 0031-3203 page 169, right-hand column, line 26 -page 170, left-hand column, line 18	1,11

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*Z\* document member of the same patent family

Date of the actual completion of the international search

7 December 2000

Date of mailing of the international search report

13/12/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Chateau, J-P

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 00/02537

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5511153 A	23-04-1996	NONE	
EP 735512 A	02-10-1996	JP 9081746 A	28-03-1997
		JP 8331607 A	13-12-1996

# RAPPORT DE RECHERCHE INTERNATIONALE

Den 3 Internationale No  
PCT/FR 00/02537

A. CLASSEMENT DE L'OBJET DE LA DEMANDE  
CIB 7 G06T7/20

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)  
CIB 7 G06T

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

WPI Data, PAJ, INSPEC, EPO-Internal, IBM-TDB

## C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 5 511 153 A (ALI AZARBAYEJANI ET AL.) 23 avril 1996 (1996-04-23) abrégé ---	1,11
A	EP 0 735 512 A (SANYO ELECTRIC CO., LTD.) 2 octobre 1996 (1996-10-02) abrégé ---	1,11
Y	J.K. AGGARWAL ET AL.: "determining motion parameters using intensity guided range sensing" PATTERN RECOGNITION., vol. 19, no. 2, 1986, pages 169-180, XP002143346 PERGAMON PRESS INC. ELMSFORD, N.Y., US ISSN: 0031-3203 page 169, colonne de droite, ligne 26 -page 170, colonne de gauche, ligne 18 -----	1,11

☐ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

### \* Catégories spéciales de documents cités:

- \*A\* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- \*E\* document antérieur, mais publié à la date de dépôt international ou après cette date
- \*L\* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- \*O\* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- \*P\* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

\*T\* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

\*X\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

\*Y\* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

\*S\* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

7 décembre 2000

Date d'expédition du présent rapport de recherche internationale

13/12/2000

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tél. (+31-70) 340-2040, Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Chateau, J-P

**RAPPORT DE RECHERCHE INTERNATIONALE**

Renseignements relatifs aux membres de familles de brevets

Dernière internationale No

PCT/FR 00/02537

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 5511153 A	23-04-1996	AUCUN	
EP 735512 A	02-10-1996	JP 9081746 A	28-03-1997
		JP 8331607 A	13-12-1996